

SCIENTIFIC DATA

OPEN

Data Descriptor: A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations

Received: 7 October 2016

Accepted: 23 March 2017

Published: 6 June 2017

Jörg Franke^{1,2}, Stefan Brönnimann^{1,2}, Jonas Bhend³ & Yuri Brugnara^{1,2}

Climatic variations at decadal scales such as phases of accelerated warming or weak monsoons have profound effects on society and economy. Studying these variations requires insights from the past. However, most current reconstructions provide either time series or fields of regional surface climate, which limit our understanding of the underlying dynamics. Here, we present the first monthly paleo-reanalysis covering the period 1600 to 2005. Over land, instrumental temperature and surface pressure observations, temperature indices derived from historical documents and climate sensitive tree-ring measurements were assimilated into an atmospheric general circulation model ensemble using a Kalman filtering technique. This data set combines the advantage of traditional reconstruction methods of being as close as possible to observations with the advantage of climate models of being physically consistent and having 3-dimensional information about the state of the atmosphere for various variables and at all points in time. In contrast to most statistical reconstructions, centennial variability stems from the climate model and its forcings, no stationarity assumptions are made and error estimates are provided.

Design Type(s)	data integration objective • observation design
Measurement Type(s)	climate change
Technology Type(s)	computational modeling technique
Factor Type(s)	DataTypes
Sample Characteristic(s)	Earth • temperature of air • Europe • England • Brazil • Jamaica • Portugal • Japan • Greenland • pressure of air • United States of America • Scotland • Russia • Hungary • Czech Republic • Germany • Poland • Switzerland • Northern Hemisphere • Austria • Romania • Sweden • Norway • Spain • Slovakia • hydrological precipitation process • atmospheric wind

¹Institute of Geography, University of Bern, 3012 Bern, Switzerland. ²Oeschger Centre for Climate Change Research, University of Bern, 3012 Bern, Switzerland. ³Federal Office of Meteorology and Climatology, Zürich Airport, 8058 Zürich, Switzerland. Correspondence and requests for materials should be addressed to J.F. (email: franke@giub.unibe.ch).

Background & Summary

Studying past decadal climate variations requires global, comprehensive data sets that are consistent with atmospheric dynamics. Ideally, such studies would be based on 3-dimensional, global representations of atmosphere and ocean in high resolution. Until recently, paleoclimatology has produced either single time series or spatially interpolated, paleodata-based reconstructions that allow to infer information mostly about past temperature and hydroclimatic conditions¹. While paleoclimatology has learned tremendously from these data sets, they may not be suitable for studying decadal variations as they do not provide global coverage, are not comprehensive (lack important variables), and may not be physically consistent. Such physical consistency, however, can only come from climate model simulations.

In atmospheric sciences, reanalysis data sets that combine measurements with model simulations have become available in the 1990s and have revolutionized the field. Atmospheric reanalyses are the most widely used data sets in geosciences, with together tens of thousands of citations. While the Twentieth Century Reanalysis project² and the European ERA-20C data set³ have shown that conventional reanalyses approaches can be extended back in time, there is a limit as to how far back the global weather can be captured. Here we present an approach that aims at a monthly time scale and exploits, at the same time, the information from early instrumental measurements, documentary data, paleodata such as tree rings, and climate model simulations.

The backward extension of conventional reanalyses was made possible by applying the Ensemble Square Root Filtering technique (EnSRF)⁴. Here we use a similar approach, but with a half-yearly assimilation step. We argue that with this assimilation frequency, initial conditions do not matter anymore, while the boundary conditions (sea-surface temperature, external forcings, etc.) provide some predictability. Therefore, the assimilation can be performed off-line on an ensemble of transient AGCM simulations (Fig. 1).

The main advantages responsible for the success of reanalysis over statistical reconstructions techniques are: i) the full spatial and temporal (4 dimensional) coverage; ii) the possibility to draw conclusions on model variables, which are not assimilated, e.g., only assimilating sea level pressure (SLP) can lead to meaningful temperature and precipitation patterns; iii) there are no assumptions about stationarity, a problem in most other reconstruction approaches because patterns may often change over time⁵. These are the reasons why researchers now start to investigate the potential of using even sparser and noisier paleo-climatic data to create a paleo-reanalysis^{6–8}. Recently, first results from an annually resolved assimilation project were published that solely relies on proxy data, where single model years from existing simulations covering the last millennium are used as ensemble members without taking their actual model year into account⁹.

This study bridges the gap between the annually resolved analysis⁹ and analyses for the last 150 years^{2,3}. It builds upon idealized model-world experiments⁶ and is the first monthly resolved paleo-reanalysis for the period 1600–2005. We assimilate early instrumental temperature and sea level pressure data, indices from historical documents as well as tree-ring width and density information into a 30-member ensemble with an atmospheric general circulation model. In the following sections, we will demonstrate the feasibility and potential of this new approach.

Methods

Data assimilation provides a best estimate of the state of the atmosphere based on climate model physics and observations (we use ‘observations’ in the following to denote instrumental data, documentary data, and proxy data). There are errors (i.e., difference to the unknown truth) on both sides. The model boundary conditions do not fully constrain the result, but the model generates random weather consistent with the boundary conditions. Its error is expressed in the mean bias and in the spread of an ensemble of simulations. Observations also have errors, e.g., due to instrument changes, reporting errors, or non-climatic signals in proxies. Again, there are random and systematic errors. Assimilation techniques calculate the most likely state of the atmosphere given all errors in both, models and observations. Here, we explain the Ensemble Kalman Fitting assimilation method on which this study is based before we present the climate model and observations used.

Ensemble kalman fitting (EKF) method

Mathematically, data assimilation can be expressed as a minimization of errors. The following cost function (1) is minimized:

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^b)^T \mathbf{P}^b (\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \quad (1)$$

where \mathbf{x} is a vector of the true atmospheric state or a time average of the state¹⁰. \mathbf{x}^b is the background or first guess, i.e., here represented by the raw model simulations. \mathbf{P}^b is the model error covariance matrix, which we calculate from the ensemble of simulations. \mathbf{R} describes the observation error covariance matrix. \mathbf{y} are the observations and the operator \mathbf{H} extracts the observations from the model space (see Experimental design and proxy forward model). In the following \mathbf{H} denotes the Jacobian matrix of $\mathbf{H}(\mathbf{x})$.

Assuming normally distributed probabilities, this cost function can be minimized with a Kalman filter, where the best estimate of the true atmospheric state \mathbf{x} is the analysis \mathbf{x}^a given by equation (2):

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{H} \mathbf{x}^b) \quad (2)$$

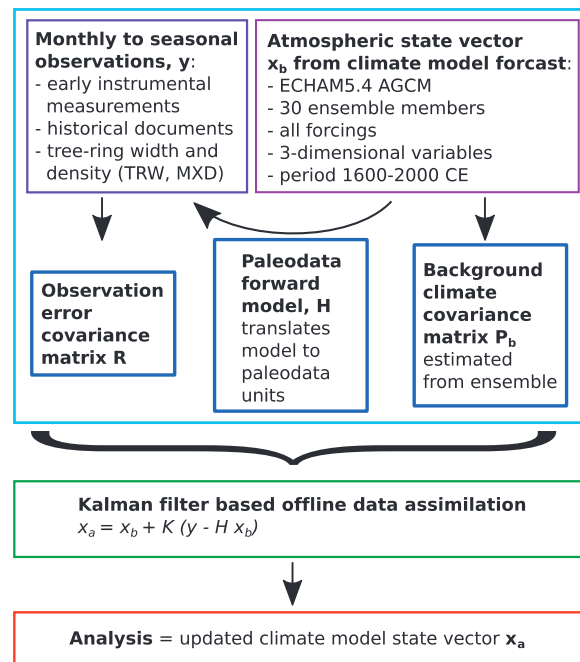


Figure 1. Schematic of the assimilation procedure.

We work with a sequential implementation of the Ensemble Kalman Filter⁴. This variation is computationally much less demanding and allows to update the ensemble members individually without explicitly updating the covariance matrices. To account for a bias occurring in the analysis covariance, we apply the ensemble square root filter¹¹. Thereby, the assimilation procedure can be split into an update of the ensemble mean (denoted by \bar{x}) and an update of the anomalies about the ensemble mean (denoted by x'):

$$\bar{x}^a = \bar{x}^b + K(y^0 - H\bar{x}^b) \quad \text{with: } y^0 = 0$$

with the Kalman gain matrices K and $K^{6,11}$:

$$K = \frac{P^b H^T}{H P^b H^T + R} \quad K^{6,11} = \frac{P^b H^T}{H P^b H^T + R} \left(I - \frac{P^b H^T}{H P^b H^T + R} H \right)$$

In numerical weather prediction, the procedure is cycled, i.e., x_a becomes the initial conditions for the next simulation step. The model thus propagates the information further into the unobserved phase space. In a paleo-climatic setting, this is not the case. The initial state is not well defined by the sparse, noisy data available in the past and the model has no skill in predicting the subsequent month from initial conditions. Nevertheless, the model simulations have skill originating from the boundary conditions (see 'data' section). Since initial conditions do not matter, a cycle is not required. This allows us to first run the entire simulation and assimilate the data afterward (off-line assimilation) instead of assimilating data after each calculation time step and then continuing the simulation (on-line assimilation)¹². We call this method, which is no data assimilation in the traditional way, Ensemble Kalman Fitting (EKF)⁶. The new reanalysis is called EKF400 because it covers the past 400 years.

Covariance localization

Because of our finite and small ensemble size of 30 simulations we have to deal with spurious correlations in the background error covariance matrix P^b . These lead to random, unphysical updates and an 'over-correction' of the analysis because very distant and in reality uncorrelated locations correlate by chance in the model world. To prevent this effect, we define a function to set a cut-off distance beyond which no update takes place following^{6,13}:

$$P_{ij}^b = \frac{1}{n_{\text{ens}} - 1} \sum_{k=1}^{n_{\text{ens}}} x_{ik}^b x_{jk}^b \exp \left(-\frac{|d_i - d_j|}{2L^2} \right)$$

with n_{ens} different ensemble members. The absolute value of $d_i - d_j$ is the distance in km between grid box i and grid box j , and L is the cut-off distance. We estimate L in the model world by calculating

Variable	Decorrelation distance (km)
Temperature (2 m)	1,500
Temperature (500 hPa)	1,500
Sea level pressure	2,700
Precipitation	450
Geopotential height (500 hPa)	2,700
Geopotential height (100 hPa)	3,750
Wind (850 hPa, longitudinal component)	1,800
Wind (850 hPa, latitudinal component)	1,800
Wind (200 hPa, longitudinal component)	1,500
Wind (200 hPa, latitudinal component)	1,500
Vertical motion (Omega, 500 hPa)	750

Table 1. Variables in the state vector and their decorrelation distances.

decorrelation distances for each variable in the state vector. This distinction is necessary as temperature for example exhibits correlations over much larger regions than precipitation (Table 1).

The assimilation step is 6-months, i.e. we assimilate April-to-September and October-to-March seasons, which are the growing seasons influencing seasonal tree-ring measurements in the northern and southern hemisphere, respectively. However, we keep monthly data for all six months in the state vector. In this way, we can assimilate monthly instrumental and documentary observations as well as seasonal tree ring-based data and achieve a final analysis with monthly resolution. Combining six months into one state vector has been successfully done for assimilating total column ozone data into chemistry-climate model simulations using the same EKF method¹⁴. Our state vector thus has a length of $n = 304,164$.

Experimental design and proxy forward model

While R and P^b are suited to measure the random part of the error, systematic biases are more difficult to treat. The model may prefer a colder state over a given region, and as soon as observations become available, they pull the model towards them, leading to a step change. Biases may also occur in the boundary conditions. Likewise paleodata often suffer from uncertain multi-decadal to centennial variability¹⁵. We solve this problem by assimilating anomalies for 70-year periods around the current year, i.e., model anomalies are corrected by the assimilation of observation anomalies. As a result, variability on scales longer than 70-yr in the analysis purely stems from the model, whereas seasonal to decadal variability is influenced by observational anomalies. To assess whether the model produces realistic multi-decadal to centennial variability, we compared CCC400 global mean land temperature with instrumental global mean temperatures from the CRUTEM4 (ref. 16) and found good agreement (Fig. 2). Further, we produced a simple reconstruction of global mean land temperature from fitting 400-year time series of external forcings (CO_2 , solar irradiance, tropospheric aerosols and stratospheric optical depth) to CRUTEM4 in a multiple regression approach with an autoregressive term (Fig. 2). CCC400 is in very good agreement with the resulting curve. It is thus reasonable to assume that CCC400 realistically reproduces variability on time scales longer than 70 years.

Assimilation entails an operator H that mimics the observations in model space, called observation operator or forward operator. For instrumental data, this means extracting the corresponding grid cell and variable from the model state vector. In the case of documentary indices, which are given in standard deviations, additional standardising (in the 70-yr window) is required. Forward operators for the paleodata tree-ring width (TRW) and maximum latewood density (MXD) are more complex. Tree rings are often influenced by more than one factor, e.g., temperature and humidity¹⁷. Proxy forward models to simulate tree-ring width such as VS or VS-Lite¹⁸ are available and could in principle be used. We decided for this study to choose a simpler approach. We used a multiple regression approach that is informed by VS-Lite. First, all TRW series were modelled with VS-lite¹⁹. From this, the relevant variables were identified as temperature during 4 months JJAS (in the northern hemisphere) and precipitation during the 3 months AMJ. The TRW series were then calibrated against corresponding CRU TS3.10 data in a regression approach. For tree-ring density we use multiple regression only for monthly temperature during the growing season as density is hardly affected by precipitation^{20–23}. Regression coefficients are calculated for a common period of overlapping paleodata and instrumental data (1901–1960).

Although the reanalysis is monthly, for simplicity we mostly present semi-annual results averaged for boreal summer (AMJJAS) and winter (ONDJFM).

Error/Uncertainty estimations

To exclude outliers from being assimilated, we first apply a variance screening by estimating local variance in the CCC400 simulations. Then, we check if the anomalies of the early instrumental data at the

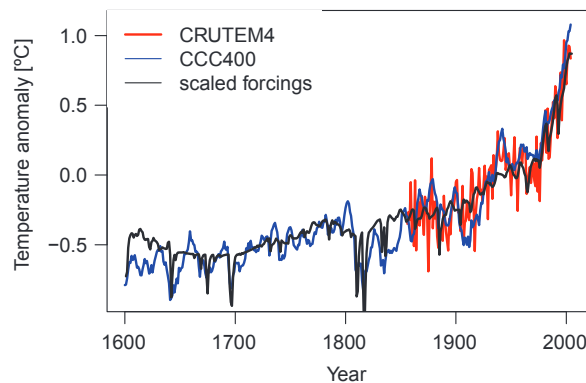


Figure 2. Global mean land temperature anomalies and scaled forcings. The agreement between instrumental measurements and simulations indicates that the model scales the external forcings to a realistic multi-decadal to centennial global mean temperature evolution: instrumental CRUTEM4 data set (red), the CCC400 simulations (blue) and the CCC400 forcings regressed to CRUTEM4 (black).

current time step is more than four standard deviations from the mean. If this is a case, we assume it to be an outlier and exclude it from the assimilation. This is justified because we do not expect extreme events such as hurricanes causing such strong deviations in monthly averages.

Estimating a specific error for each individual record that gets assimilated is practically impossible because information about data quality exists only for a few instrumental records that have been homogenized recently. For early instrumental data, all records that pass the screening get a common plausible error assigned (1K for temperature and 3 hPa for SLP). Error contributions come from various sources. One major contribution is usually assessed by data homogenization projects and includes errors such as shading of thermometers or uncertainties due to human reading of the instrument. Altogether, we estimate this part to be around 0.5 K based on expert knowledge. A second error source is stations not being representative for a grid box average. The latter error has been estimated by MeteoSwiss for one of their gridded data sets to be around 0.5 K in the non-alpine regions of Switzerland²⁴. Hence, we believe an error of around 1 K to be a realistic first guess estimate.

Instrumental measurements can be assumed to be more precise than interpretations of historical documents but the documentary data is often only available as anomalies in units of standard deviations. Thus, we set the documentary data error to 0.5 standard deviations. Note that only documentary temperature information is assimilated.

For the tree-ring data we use the variance of the multiple regression residuals as an error estimate. i.e., if the tree-rings contain a strong climatic signal, the multiple-regression model will create a good fit to instrumental data and consequently the residuals will be small.

Data Records

Model simulations

In this study, we use the atmospheric general circulation model simulations⁶, which we call ‘CCC400’. This is a 30-member ensemble of ECHAM5.4 (ref. 25) simulations at a triangular spectral truncation of T63 and 31 levels in the vertical covering the period 1600 to 2005.

The simulations are driven by reconstructed and observed monthly variations in aerosol optical depth caused by volcanic eruptions²⁶ and variations in tropospheric aerosols²⁷, total solar irradiance²⁸, greenhouse gases²⁹, land surface conditions³⁰, sea ice cover and sea surface temperatures (SST). The reconstructed annual SSTs³¹ ($SST_{\text{year mean}}$) are superimposed with a seasonal anomaly (SST_{seas}) and ENSO dependent anomalies (SST_{ENSO}), both based on HadISST1.1 (ref. 32), with SST_{ENSO} being estimated by least squares regression on the reconstructed annual El Niño 3.4 index³³ with lag 0 and lag -1:

$$SST \approx \frac{1}{4} SST_{\text{year mean}} + \frac{3}{4} SST_{\text{seas}} + \frac{1}{4} SST_{\text{ENSO}} \quad \text{Eq. 1}$$

For lack of a skillful sea-ice reconstruction, we use the sea-ice climatology from the HadISST1.1 data set before 1870. This is expected to result in an underestimation of high latitude climate variability in the simulations.

Unfortunately, land-surface changes have not been correctly included, which caused an overestimation of its effects on albedo, surface roughness, vegetation variables and field capacity. We assessed the bias with one corrected simulation that could be conducted additionally. Thirty-year mean temperatures over extra-tropical northern hemisphere (20–90°N) land areas are on average 0.28 K (−0.5 to +1.2 K locally) warmer at the beginning of the 17th century and are on average 0.31 K (−1 to +2.4 K locally) warmer in

Name	Type	Period	Temporal resolution	Area	Spatial resolution	Variables	Reference
CCC400	ECHAM5.4 all forcing simulation	1600–2005	6-h	Global	2° × 2°	All ECHAM output variables	Bhend et al. ⁶
EKF400	Assimilation (30 ensemble members)	1600–2005	Monthly	Global	4° × 4°	Temperature, precipitation, sea level pressure, wind, geopotential height, vertical motion	This study

Table 2. Data sets used and generated in this study.

the late 20th century than the corrected simulation. Other variables such as pressure, wind and precipitation are also slightly affected. Since the model is debiased in the pre-processing scheme using a 70-yr window, effect of the land surface specification on the assimilation will be small (see ‘Experimental design’) for monthly to seasonal averages, which is the target of this study. Additionally, only around Antarctica, excessively high and unrealistic wind speeds are reached, which cause wave-like patterns poleward of 60° S. Therefore, we exclude any interpretation of the latitudes south of 60° S. To keep the assimilation computationally tractable we use every second grid box and only keep the variables seen in Table 2 in the state vector. Each variable has 4,608 grid boxes over land and ocean.

Observations

We assimilate monthly instrumental measurements of temperature and sea level pressure, monthly temperature information from documentary sources and half-yearly tree-ring paleodata: tree-ring width (TRW) and maximum latewood density (MXD). We always use all information available (m records) except for instrumental data where we freeze the network at the year 1880 state, i.e., we exclude all series that start after 1880. Our main goal is to construct a paleo-reanalysis, for the 20th century there is already a higher resolved product available². Thus, we use the 1880 data network during the 20th century mainly to evaluate the quality of our reanalysis in the past. The first instrumental series start in the year 1659, the last documentary series end in 1853 and all TRW and MXD records end in 1960 (Fig. 3).

Instrumental temperature series include the GHCN collection³⁴ the collection of the Climate Research Unit at the University of East Anglia³⁵, historical instrumental time series of the Greater Alpine Region (HISTALP)^{36,37}, as well as temperature series from Central England³⁸, Rio de Janeiro³⁹, Funchal Madeira^{40,41}, Nagasaki, Tokyo and Yokohama⁴², Illussat, Nuuk and Qaqortoq⁴³ and 18th Century temperature measurements from Jamaica⁴⁴. Instrumental SLP mainly comes from GHCN⁴⁵ and HISTALP, too. Additionally, we incorporated a collection of European SLP measurements⁴⁶, series from Salem Massachusetts⁴⁷, GordonCastle⁴⁸, Nagasaki and Tokyo⁴² and Nyzhny Tagil⁴⁹. We use monthly resolved documentary temperature information for the Carpathian basin⁵⁰, the Czech Republic⁵¹, Germany⁵², Poland⁵³ and Switzerland⁵⁴. We use 35 climate sensitive TRW records^{8,19}. MXD records mainly stem from a previous collection⁵⁵. Additionally, we use density records from the Austrian Alps⁵⁶, Swiss Alps⁵⁷, Carpathians⁵⁸, Jeamtland⁵⁹, Fjorfordalen²³, Pyrenees⁶⁰, Tatra⁶¹ and Tornetrask²¹.

In very few places in Europe and the United States of America we must deal with multiple time series that are located in the same model grid box. In these cases, we prefer instrumental series to documentary series to paleodata, i.e., if there is an instrumental measurement in a grid box, we discard the paleodata. If more series of the same type exist, we calculate a grid box average. All instrumental data are in the form of anomalies (see section ‘experimental design’). Therefore, biases, e.g., due to different elevations above sea level, can be neglected.

Validation data sets

To assess the skill of our analysis we compare it with several observational data sets for the 20th century. We chose the gridded temperature, precipitation and SLP data sets CRU TS 3.10 (ref. 62). To assess uncertainties in the gridded instrumental data sets, mostly due to measurement errors and interpolation, we consult the CRUTEM ensemble of instrumental fields¹⁶. Before 1900, we conduct a leave-one-out validation and compare our assimilation product to the spatial reconstruction^{46,63,64} for temperature, precipitation and SLP in Europe, respectively. Furthermore, we use several proxy based reconstructions of northern hemisphere land temperature^{65,66}, calculated atmospheric circulation indices⁶⁷ and reconstructed indices^{68–72} as well as the 20th century reanalysis² version 2c (20CR) and European ReAnalysis ERA-40 (ref. 73).

Technical Validation

Estimating the skill of EKF400 is challenging because the network of assimilated data and hence the skill varies with each time step. Moreover, the goal would be to assimilate all data available, thus no completely independent data would be left for validation. For this reason, we present a set of skill measures: comparisons with 20th century instrumental observations, time series of large-scale average temperature from observations and paleodata. Finally, we present a case study for the 1815 Tambora eruption, which caused severe climatic conditions 200 years ago. We present also completely independent results for precipitation and selected atmospheric indices.

Analysis skill is assessed with the following measures: First we calculate the Pearson correlation coefficient to evaluate the covariability of data sets. Second, we use the Reduction of Error (RE) statistic⁷⁴,

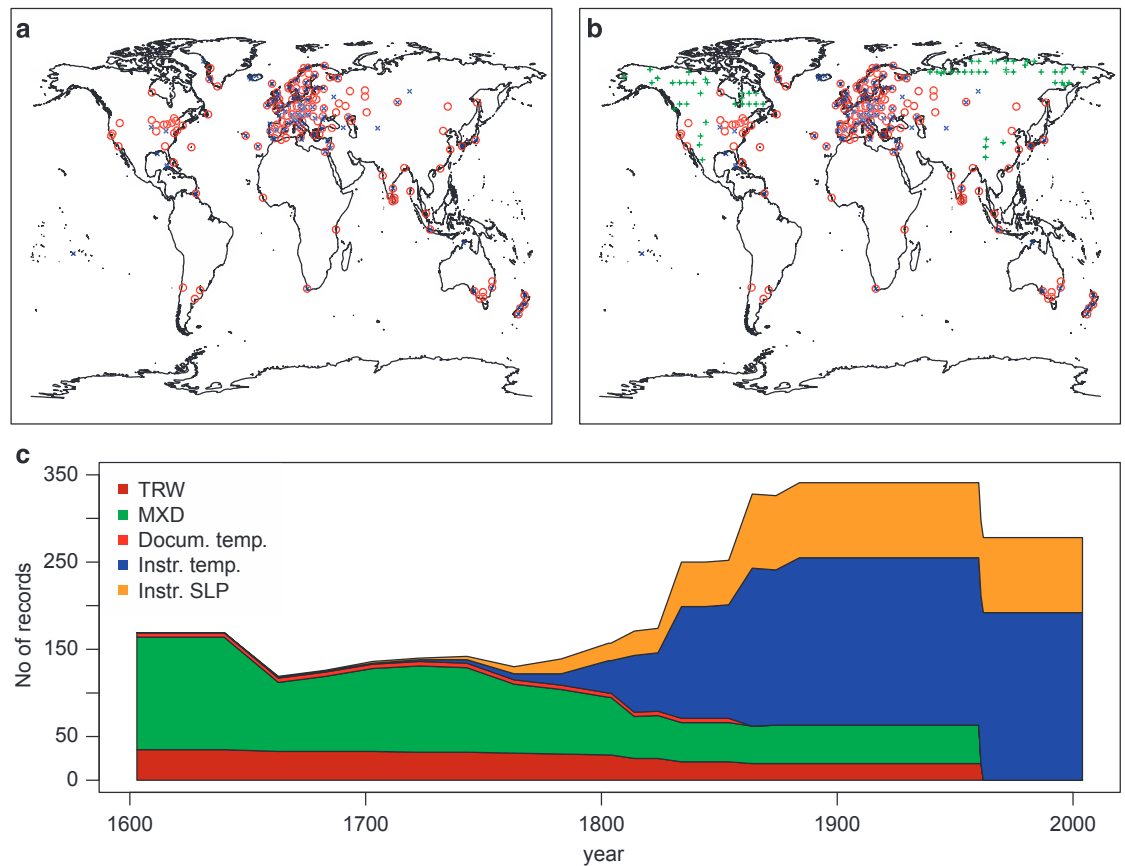


Figure 3. Spatial and temporal overview of assimilated data. Locations with instrumental temperature measurements are highlighted with red circles, sea level pressure stations with blue crosses and tree-ring width/density locations with green 'plus' symbol in October to March (a) and April to September (b) of the year 1880. The time dependent number of input data is shown in (c) separated by data type and variable.

also known as the mean squared error skill score⁷⁵. RE is calculated with the following equation:

$$RE = 1 - \frac{\sum_i (x_i^a - x_i^{ref})^2}{\sum_i (x_i^b - x_i^{ref})^2} \quad (3)$$

where x^a is the analysis EKF400, x^b is a 'no knowledge prediction' and x^{ref} is the reference, in our case the CRU TS 3.10 instrumental data. As no knowledge prediction we choose CCC400 in which case values above 0 indicate that EKF400 is closer to CRU TS3.10 than CCC400 and hence the assimilation step has added information. Note that RE uses squared absolute deviations, i.e., RE can be negative even if correlations are strongly positive.

The ensemble-based error estimate is a major advantage of our assimilation in comparison with previous statistical reconstruction approaches. We highlight how the standard deviation, often called 'spread', of the ensemble simulations is reduced by the assimilation (Fig. 4). This alone, however, is not enough to judge the reliability of the system. Reliability can be evaluated with the spread-error ratio (SPR2ERR):

$$SPR2ERR = \frac{\frac{1}{n_{ens}} \sum_i s_{ens,i}^2}{\frac{1}{n_{ens}} \sum_i MSE_i} \quad (4)$$

Here, the temporal mean of the ensemble variance $\overline{s_{ens}^2}$ is divided by the mean squared error of the observations from the ensemble mean (MSE). The intra ensemble sample variance is inflated to take the finite ensemble size (n_{ens}) into account⁷⁵. Values of SPR2ERR close to 1 indicate an appropriate ensemble spread and consequently the observations are consistent with the analysis and analysis error estimate. Values of SPR2ERR smaller than 1 indicate overconfidence with insufficient ensemble spread. In our assimilation, overconfidence can be caused by a variety of factors including insufficient

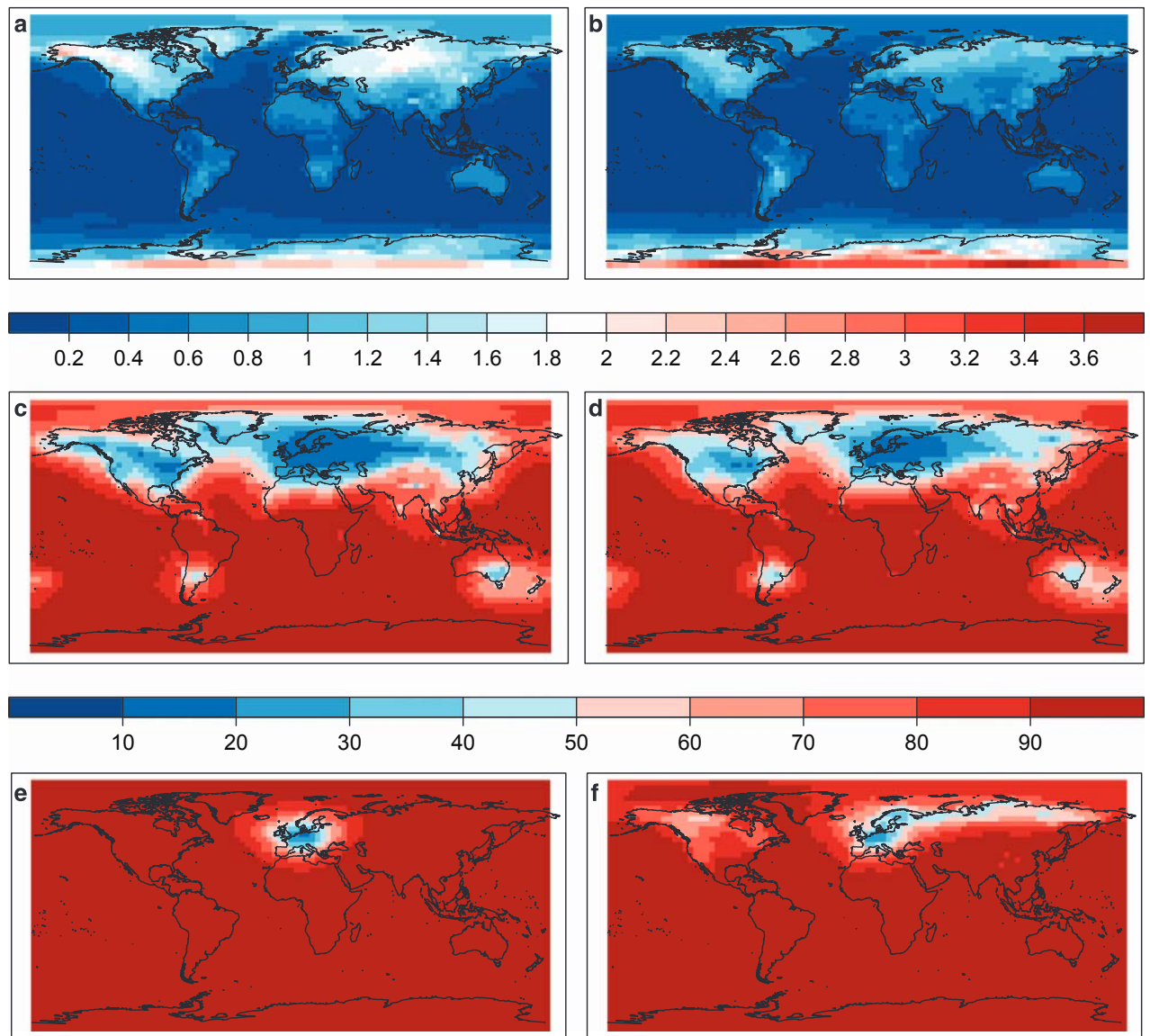


Figure 4. The data assimilation reduces the temperature spread of the 30 CCC400 ensemble members. The original CCC400 spread is shown for October to March (a) and April to September (b). The remaining percentage of ensemble spread after the data assimilation can be seen during the verification period 1902–2001 in (c, d) and for 1651–1750 when mostly tree-ring measurements are assimilated in (e, f). The October to March season is shown in the left column (a, c, e) and the April to September season in the right column (b, d, f).

localization, underestimation of observation errors, correlation of observation errors, or insufficient spread in CCC400. Values of SPR2ERR larger than 1 on the other hand indicate overdispersion.

This SPR2ERR definition assumes error-free observations. Errors in observations can be treated as additive noise^{76,77}. Our error adjusted MSE_a is given by:

$$MSE_a = \frac{1}{4} MSE - \frac{SD_{obs}^2}{811P}$$

where SD_{obs}^2 is the variance of the CRUTEM4 ensemble estimate of gridded instrumental temperature observations¹⁶.

To evaluate the reanalysis skill before the 20th century and its variation over time, we conduct a leave-one-out (LOO) validation for the period from 1760 to 1880. We systematically discard instrumental temperature observations of a single grid box and compare the resulting analysis with the independent, discarded measurements. This gives us two important pieces of information. First, at each grid cell with measurements, we can calculate the correlation and MSE with independent data. Second, with the ensemble spread of the LOO analysis and the MSE we can calculate the SPR2ERR further back in time. In contrast to the 20th century, when information was spatially nearly complete, in the earlier period we

only have information at a few locations with observations. Thus, the temporal evolution of SPR2ERR change only represents this area with observations.

Reanalysis skill

First, we assess the skill with a focus on the 100yr period 1902–2001 CE, when the overlap with instrumental observations allows for comparisons. We evaluate EKF400 with CRU TS3.10 as a reference.

Within the CCC400 ensemble we find the largest temperature spread in continental regions and high latitudes during the winter season of each hemisphere (Fig. 4, top). Data assimilation reduces the ensemble spread with respect to measurement uncertainty where the covariance matrix indicates spatial relationships within the localization distance. In regions with considerable spread and with a dense network such as Europe, the spread can be reduced to 10–20% of the original CCC400 spread whereas it remains unchanged at nearly 100% distant from any observation, e.g., in central Africa (Fig. 4, middle). Looking back into the period 1651–1750 of EKF400, when only European documentary information is assimilated in October to March, the spread reduction is limited to Europe. During the April to September season, when tree-ring data is available in most extra-tropical land regions of the northern hemisphere spread can additionally be reduced in northern Russia and northern North America (Fig. 4, bottom).

The SPR2ERR for the 20th century suggests that overfitting (too much reduction of model spread without comparable reduction of error) occurs in the 20th century in large parts of North America and Eurasia (Fig. 5, top). If uncertainties in the observations are considered, we only find clear overfitting in Europe and slight overfitting in the United States of America (Fig. 5, bottom). This could have two potential reasons. First, the spread of CCC400 is already too small because of the relatively small ensemble size of 30 members. This can be ruled out by the spread-error analysis for CCC400 (not shown) that indicates too little spread at very few coastal grid boxes in the tropics, which are most likely a result of the SST forcings. Second, the observation error is too small or correlated. Unfortunately, there is no quality information for most of the records and we have only been able to assign a best guess error from the literature to all series. Additional analysis suggests that representativity errors may be larger than prescribed as our grid boxes span a larger area than in the study of MeteoSwiss²⁴. The challenge of assigning adequate observation errors requires further investigation in the future. In contrast to Europe and North America, spread-error ratios for central South America and central Africa indicate that the model spread could not be reduced sufficiently.

SPR2ERR from the LOO experiments confirm the results from the 20th century. When hardly any instrumental temperature data is assimilated such as from 1760 to 1770, median SPR2ERR values of 1.5 indicate overdispersion (Fig. 6e). From 1780 onward the SPR2ERR decreases with median values reaching optimal ratios of one around 1800 and asymptotically approaching values of 0.5 indicating overconfidence at the end of the 19th century. The main reason for overconfidence is a reduction in ensemble spread over central and eastern Europe without a significant reduction in MSE. The fact that assimilating few measurements leads to a desired SPR2ERR ratio of one but a denser data network to ratios below one suggests that measurement errors are currently underestimated. Hence, future methodological improvements will focus on finding more data in sparse regions and further back in the past as well as improvements in assigning measurement errors, localisation distances and observation error correlation.

Ensemble mean simulated temperature and observed temperatures are positively correlated already before data assimilation has been applied, in particular in coastal regions. The main reason is the model forcing and here specifically the prescribed sea surface temperatures. Data assimilation, however, greatly enhances positive correlation between EKF400 and instrumental CRU TS3.10 data, as seen in the correlation differences before and after the assimilation (Fig. 7a,b). Only in limited regions of central South America, central Africa and the Middle East, which are far away from any observation, we find no improvement in correlation. Although we do not assimilate any precipitation data, EKF400 has much higher correlations with instrumental precipitations measurements than CCC400 (Fig. 7c,d) due to covariances between temperature/SLP and precipitation. This increase in correlation is strongest in Europe and the United States, where most data gets assimilated, but improvements can also be seen in southern South America Southwest Australia and New Zealand. LOO experiments confirm the increase in correlation between independent validation data and the analysis in almost all location for the period 1761–1880. Highest correlation coefficients are reached in Europe and the eastern United States of America where the observation network is densest and nearby observations can substitute the observation that has been left out (Fig. 6a,b).

If the RE is used to evaluate improvements of EKF400 ensemble mean over CCC400 ensemble mean, we similarly find positive skill in temperature for most of the northern hemisphere apart from central/northern Greenland and eastern Russia (Fig. 8a,b). In these regions, however, negative values may actually mean an improvement, because the instrumental data set that is used for validation has too little variability at grid cells far away from any station where values were either interpolated or long term averages have been filled in. This is the case, for example, over central Greenland, where CRU TS3.10 has basically no variability. In theory, spurious correlation in the covariance matrix, which is based on 30 ensemble members only, could lead to such errors in the analysis but in general the localization prevents such corrections. In regions far away from any assimilated data, e.g., in Africa, RE values indicate zero

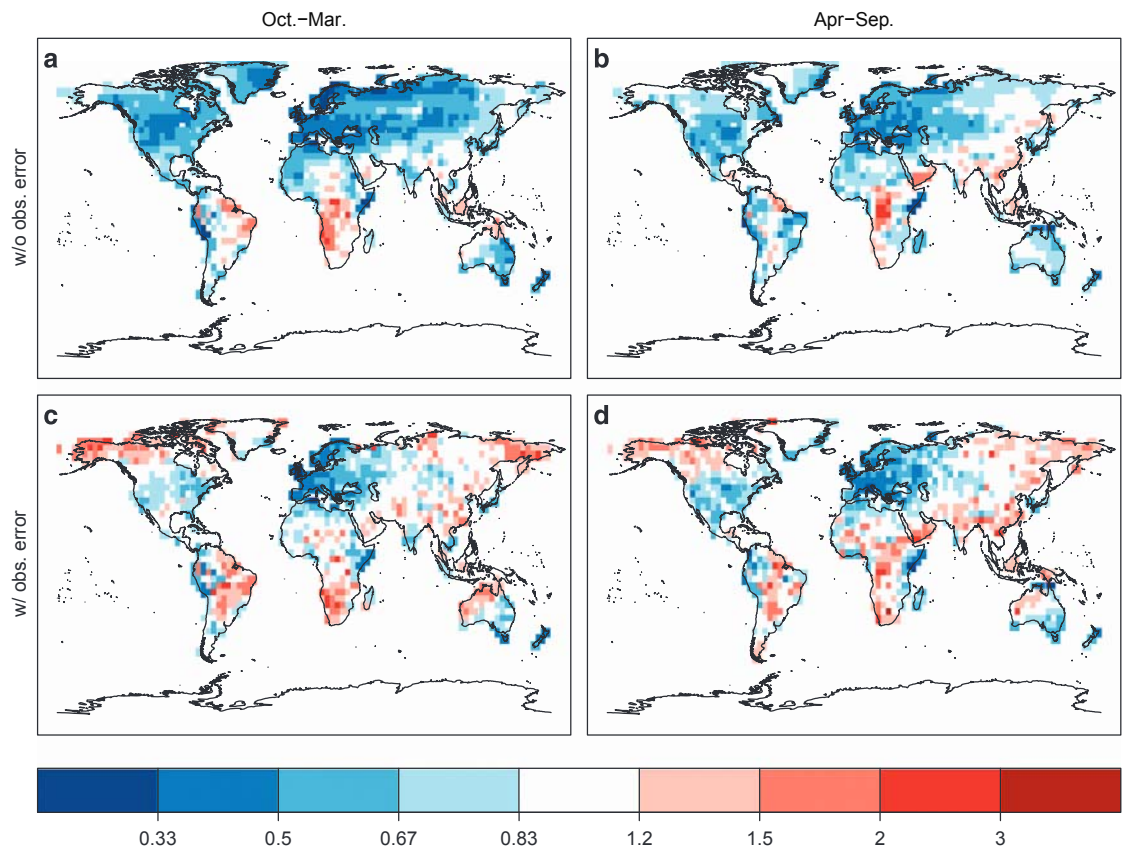


Figure 5. The spread to error ratio of EKF400 during the 20th century indicates where and when the analysis is overdispersed or overconfident. The maps at the top (a, b) result from the assumption of perfect observation whereas errors in observations are accounted for in c, d.

skill, i.e., EKF400 does not differ from CCC400. RE values in the LOO experiments for the years 1761–1880 are globally around zero (not shown). The reason is not that the LOO analysis ensemble mean would be equal to the CCC400 ensemble mean, LOO has clearly enhanced variance. Although the amount of variance itself is closer to the variance of the observations, it causes the low RE values because both overestimation and underestimation of the observed value get punished in the RE statistic. That is why the low variability CCC400 ensemble mean can achieve the same RE values than the LOO ensemble mean although LOO variability and correlation indicate that the LOO analysis is superior to CCC400.

The RE values for precipitation are negative in some regions of Europe, Asia and North America, indicating that the assimilation of temperature and SLP is not sufficient to lead to an improvement in absolute precipitation amounts. Additional experiments that assimilated precipitation measurements did not lead to better skill in EKF400 precipitation, either. Precipitation is locally very variable and not spatially correlated over larger regions. Due to regional covariability between temperature/SLP and precipitation, EKF400 mostly has the correct sign of the precipitation anomaly (Fig. 7) but not the correct amount (Fig. 8).

How the assimilation works and what differences in correlation and RE mean, can be best understood in an example. The monthly average temperature at a grid box in Norway (62° N, 11° E) in the CCC400 ensemble mean shows too little variability, has a correct annual cycle, but does not correlate with instrumental data, if the annual cycle is removed, because of internal variability in the model (Fig. 9). The assimilation of temperature and pressure measurements leads to a nearly perfect match between instrumental temperature and EKF400. For precipitation, the CCC400 ensemble also has too little variability. Here the assimilation also causes an increase in variability, mostly pointed in the correct direction. However, this increase is often too large, which causes the absolute values of EKF400 being further away from instrumental data than CCC400. For SLP the assimilation works nearly as well as for temperature. We also see the overconfidence suggested by the spread-error ratio in Fig. 5, because some of the temperature and SLP values lie outside the narrow band of EKF400 spread.

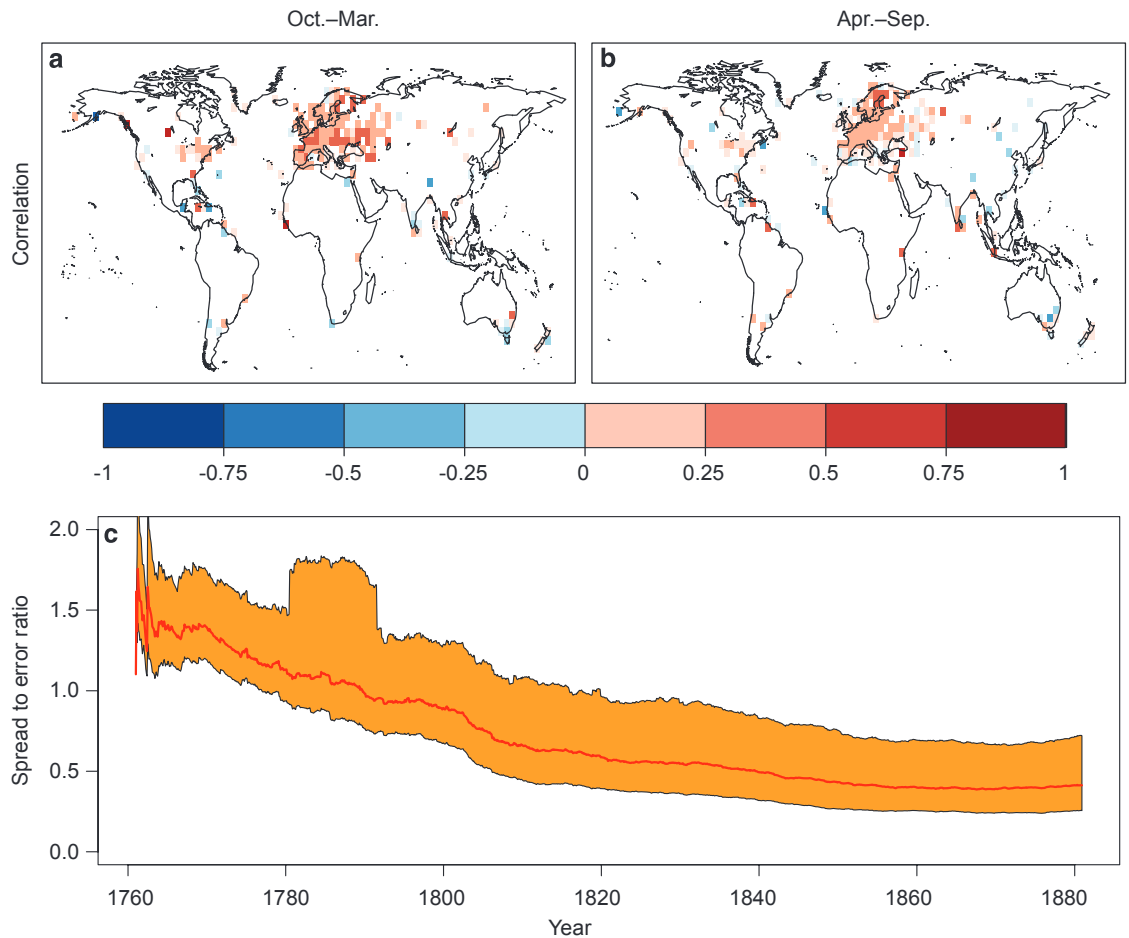


Figure 6. The EKF400 skill in the period 1761–1880 is assessed with a leave-one-out validation. The maps (a, b) show the Pearson correlation coefficients between each excluded and thus independent observation and the analysis with all remaining observations assimilated. The October to March season is shown on the left (a), April to September on the right (b). The time series in (c) highlights the evolution of the median spread to error ratio (red) and the 0.25 to 0.75 quartile range (orange) of the leave-one-out experiment as 11-year moving averages.

Northern hemisphere temperature variability

One important aspect of paleoclimatology has long been the reconstruction of hemispheric to global temperature variability to place the current global warming into a longer-term perspective. Although a lot of progress has been made both in methodology and the amount of paleodata records for single locations, substantial uncertainties remain until today. These are often related to the selected set of records, the applied statistical method, seasonality and their regional representativity⁷⁸.

Large-scale temperature averages vary from annual to centennial and longer time scales due to internal weather processes and external forcings. In Fig. 10 extratropical northern hemisphere (ENH, 20–90° N) summer (April to September) land temperature variability has been plotted from 1600 to 2005 CE. because that is what northern hemisphere temperature limited tree-rings record, in which play the dominant role in most previous reconstructions. Additionally, this is the region and season, in which EKF400 is expected to have most skill, especially in the early part when mostly tree-ring data are assimilated. The light red band in Fig. 10 highlights the spread of the EKF400 ensemble. It has a spread of roughly 0.7 K before 1800 when mainly tree-ring data is assimilated and a spread decreasing to ~0.3 K after 1800 with the increase in the number of instrumental measurements. The co-evolution of the EKF400 ensemble mean, recent proxy-based reconstructions and instrumental data sets indicate clear improvements due to the assimilation (Table 3). To remove the influence of long term trends or multi-decadal variability on the correlation coefficient, we calculate correlations not only based on the original series but additionally on a 11-year high pass filtered version, in which low frequency variability has been removed. The results highlight that CCC400 is highly correlated with the reference series because of common multi-decadal changes prescribed by the forcings. On sub-decadal time scales

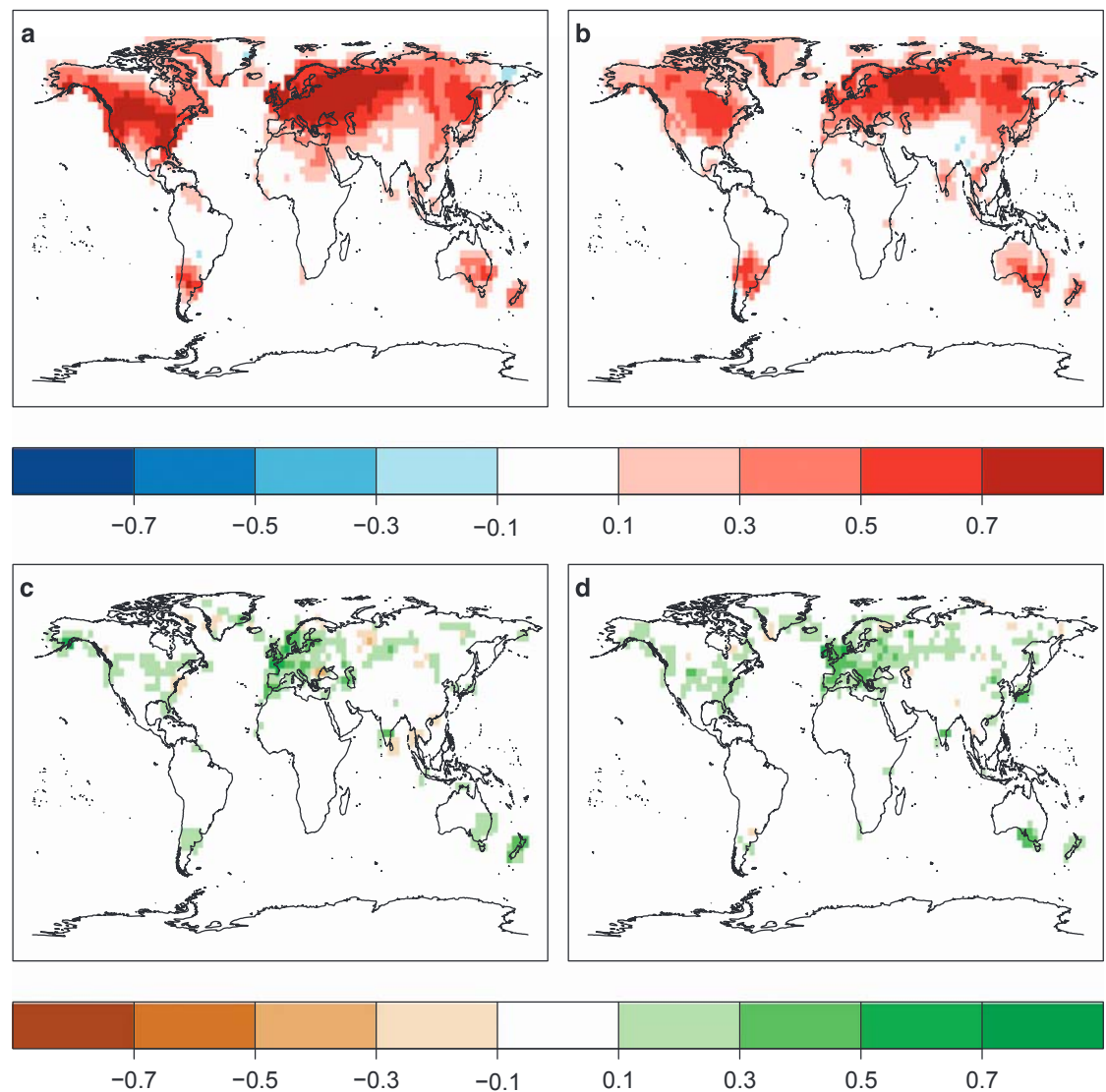


Figure 7. The increase in correlation highlights the improvements due to the assimilation procedure. The maps show differences of correlations between gridded instrumental CRU TS 3.10 data and CCC400 and EKF400. Temperature is presented on top (a, b), precipitation at the bottom (c, d). The October to March season is shown on the left (a, c), April to September on the right (b, d).

EKF400 correlates better with the instrumental data and paleodata reconstructions. The skill in EKF400 is also visible in the clearly positive RE values with a maximum of 0.79 if compared to CRUTEM4 (ref. 16). Similar to the global mean, we also find high agreement between continental scale temperature reconstructions of the PAGES 2k Consortium^{79,80}. Note, the underestimation of ENH average temperature in the last 35 years is an artefact of using 70 year average anomalies with respect to the current year. For the past 35 years, we had to shorten this period and keep the same final year in the calculation of the anomalies.

The ensemble standard deviation of northern hemisphere extra-tropical land temperature in CCC400 is 0.35/0.19 (Oct-Mar/Apr-Sep) due to internal variability in the model. EKF400 has a standard deviation of 0.32/0.18 between 1600–1800 CE. Values decrease rapidly in the first 20 years of the 19th century down to 0.13/0.09. In comparison, the CRUTEM4 ensemble suggests observation and interpolation errors in the instrumental data with values of 0.11/0.09 is very close to the level that EKF400 reaches at the early 19th century.

Case study: 1816, the year without a summer in Europe after the eruption of Mt. Tambora
To look a bit closer into the performance of EKF400 in earlier times and at the possibilities it offers for understanding the climate dynamics, we present a case study for the summer 1816 in Europe, which followed the Tambora eruption in April a year earlier and which led to the last subsistence crisis in

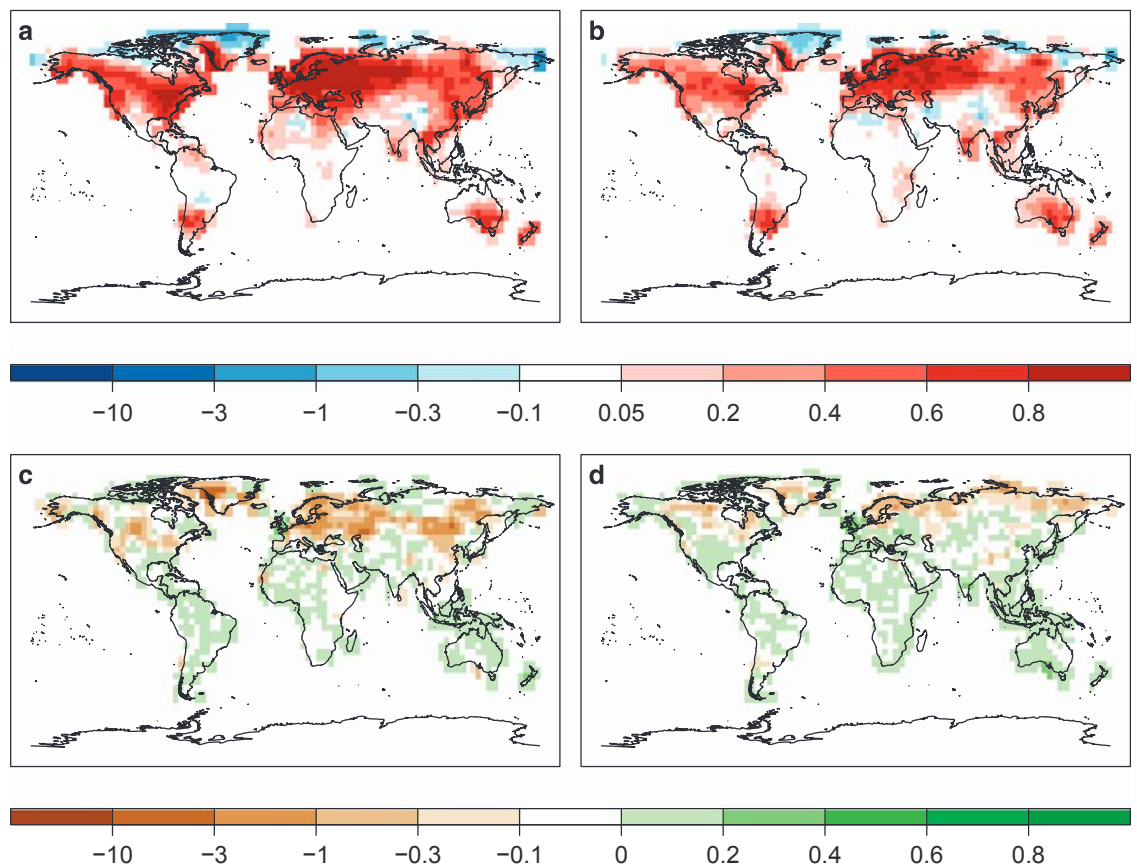


Figure 8. Positive values of the reduction of error (RE) statistic indicate that the reanalysis is closer to the validation data than the CCC400 ensemble mean. RE for temperature and the seasons October to March (a) and April to September (b) and precipitation (c, d).

central Europe⁸¹. For the case of Geneva, Switzerland, only about half of the cooling could be related to the volcanic aerosol forcing and the other half was related to more frequent westerlies, which brought rain from the Atlantic to central Europe⁸².

In the CCC400 simulations, which are forced with the reconstructed aerosols, we find a cooling of ca. 0.5 K over Europe and neither strong anomaly in SLP nor precipitation (Fig. 11a,d). The assimilation leads to a much stronger temperature anomaly of around -1.5 K in central Europe and a SLP anomaly of -2 hPa (Fig. 11b). These values in the EKF400 ensemble mean are very close to the reconstructions of temperature⁶³ and SLP⁴⁶ where anomalies of -2 K and -2 hPa have been found (Fig. 11c). Interestingly, EKF400 suggests no warming, but normal temperatures in Eastern Europe, while reconstructions indicate warming.

The latter two data sets are partly based on the same input data. This is not the case for precipitation. EKF400 shows a strong positive precipitation anomaly over central Europe (Fig. 11e), similar to reconstructed precipitation⁶⁴ (Fig. 11f). Beyond this, EKF400 also highlights the enhanced westerlies in the wind anomaly field, which have been described based on early observations⁸² (Fig. 11b).

Atmospheric indices

EKF400 allows for deriving large-scale averages or indices from its spatial fields. With the current setup and paleodata network there is skill in the North Atlantic Oscillation index (NAO) and the Pacific North American pattern (PNA) (Fig. 12), both calculated as annual averages over the months December to March. Skill in the NAO increases from near zero correlation between CCC400 and 20CR² and ERA40 (ref. 73) to correlations of around 0.8 with EKF400. This may come to less of a surprise because SLP data from Iceland and Portugal gets assimilated. However, during the period from 1600–1900 correlations with NAO reconstructions^{68,69} increase from zero in CCC400 to 0.36/0.61 in EKF400, respectively, although few SLP observations are assimilated in the 19th century and none before. Only the low frequency reconstruction by Trouet *et al.*⁷⁰ appears to disagree with all the other reconstructions including EKF400.

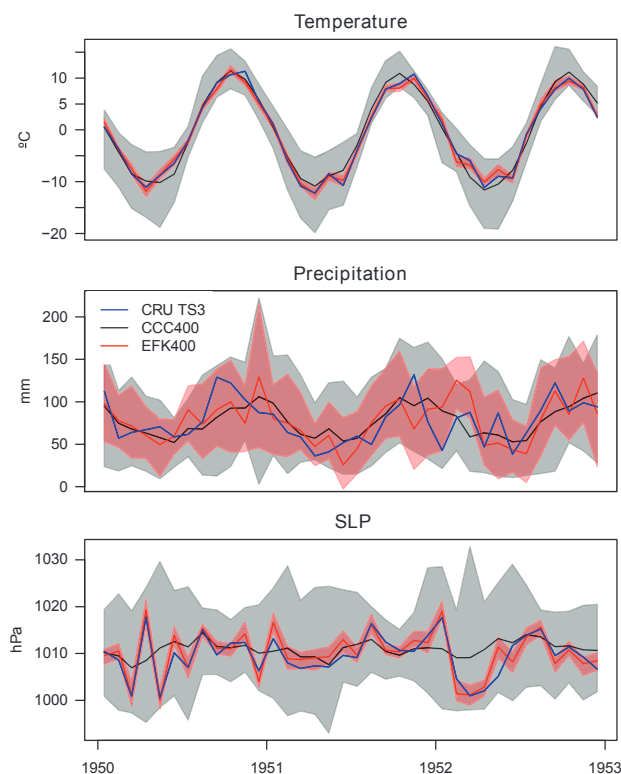


Figure 9. These examples highlight how the assimilation affects the final reanalysis. The time series are monthly averages for 3 years in Norway (62° N, 11° E) for temperature (top), precipitation (middle) and sea level pressure (bottom), all include instrumental data (blue), CCC400 ensemble mean/spread (black/grey) and EKF400 ensemble mean/spread (red/light red).

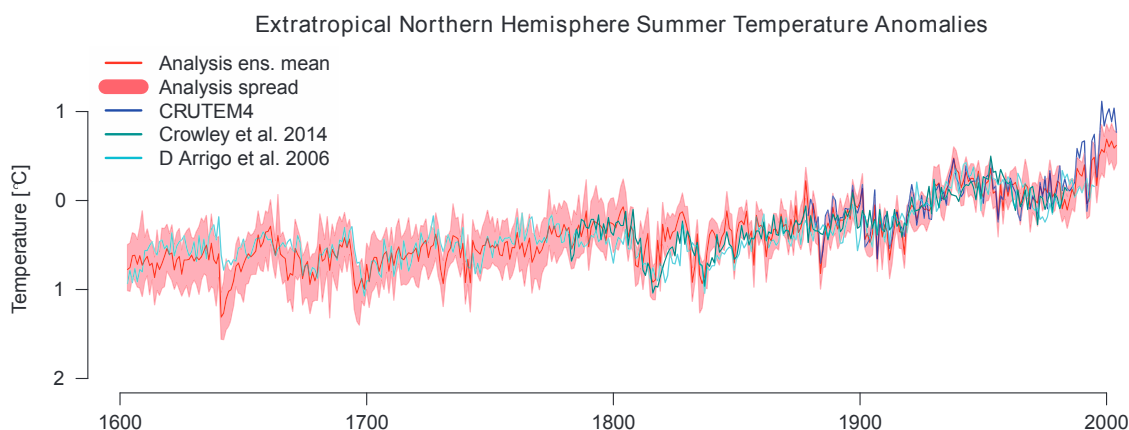


Figure 10. Northern hemisphere temperature evolution over the past centuries is a key figure in many paleoclimatic reconstructions. Here we present average extratropical northern hemisphere summer land temperature anomalies (wrt 1901–1980) 2 m above the surface EKF400 ensemble mean and members (dark and light red), CRUTEM4 gridded instrumental data (blue) and selected reconstructions (light and dark cyan).

PNA is calculated from geopotential height fields at 500hPa and skill in this region solely stems from the assimilation of a few temperature observations and the many tree-ring measurements. Here, correlations with 20CR² and ERA40 (ref. 73) in the 20th century increase from around 0 for CCC400 to ca. 0.55 for EKF400. However, we neither find any agreement between EKF400 and a previous reconstruction⁷¹ in the 20th century nor before. In pseudoproxy experiments⁶ skill has been found in

	r CRU before/after 11-year high pass filtering	r Crowley before/after 11-year high pass filtering	r D'Arrigo before/after 11-year high pass filtering	RE CRU	RE Crowley	RE D'Arrigo
CCC400	0.76/0.27	0.85/0.19	0.81/0.12	0.79	0.18	0.37
EKF400	0.85/0.81	0.51/0.37	0.48/0.26			

Table 3. Pearson correlation coefficients (r) between CCC400/EKF400 and instrumental data/selected reconstructions, all for the extra-tropical northern hemisphere land temperature mean. The first value is based on the original time series including all frequencies and after applying an 11-year high-pass filter to focus on improvements in inter-annual variability that is less influenced by external forcings. RE values indicate how much closer EKF400 is to the reference than CCC400. Always the entire overlapping period of the series is considered (CRU: 1850–2004; Crowley: 1782, 1984; D’Arrigo: 1608–1990).

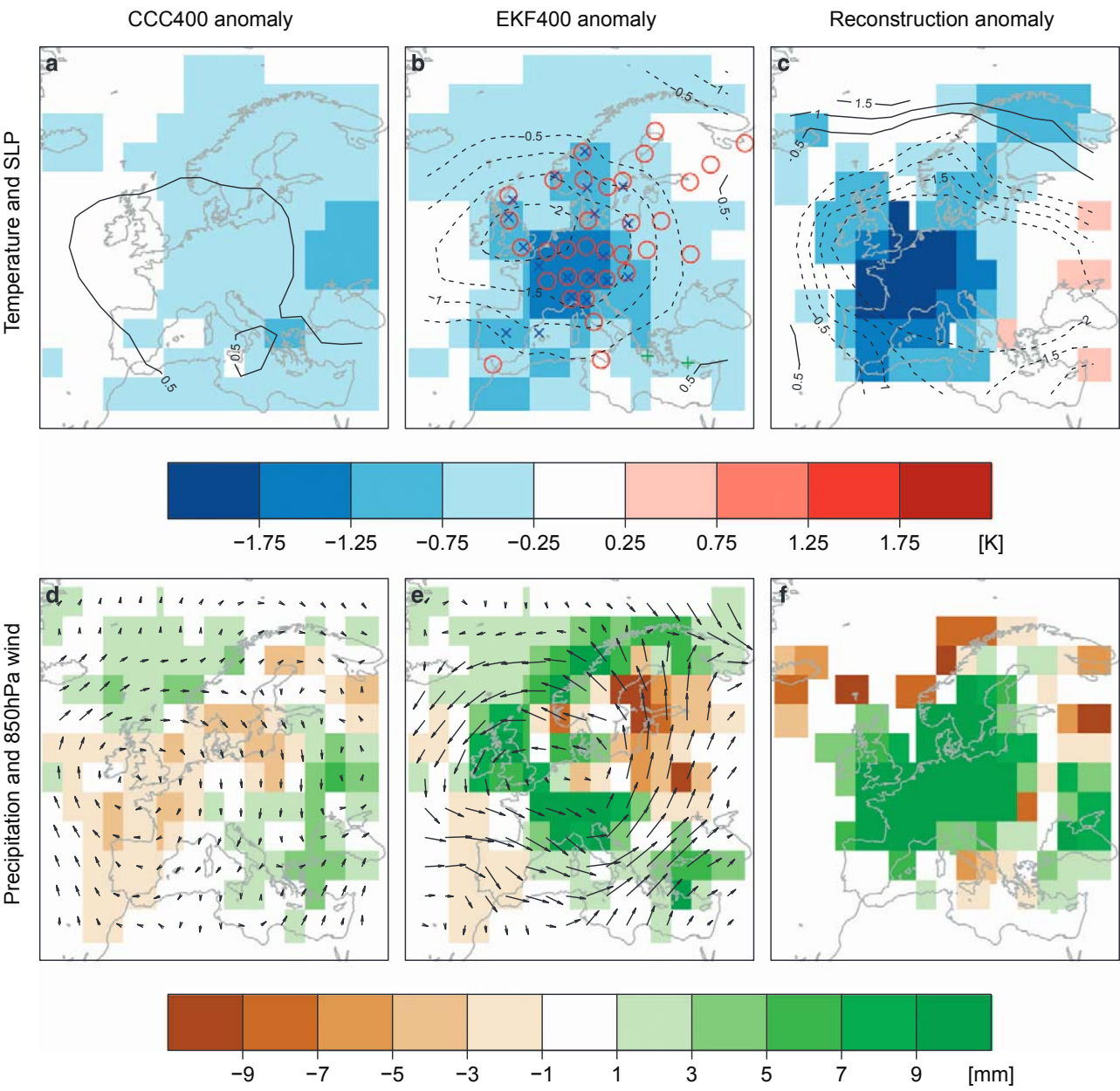


Figure 11. A great advantage of paleo-reanalysis over statistical reconstructions is the possibility the interpret the multivariate state of the atmosphere. Temperature and SLP anomalies for April to September 1816 in CCC400 (a), EKF400 (b) are compared to the reconstructions by Luterbacher *et al.* 2004 and Küttel *et al.*⁴⁶ (c), precipitation and 850 hPa wind anomalies in CCC400 (d), EKF400 (e) to the reconstruction by Pauling *et al.*⁶⁴ (f).

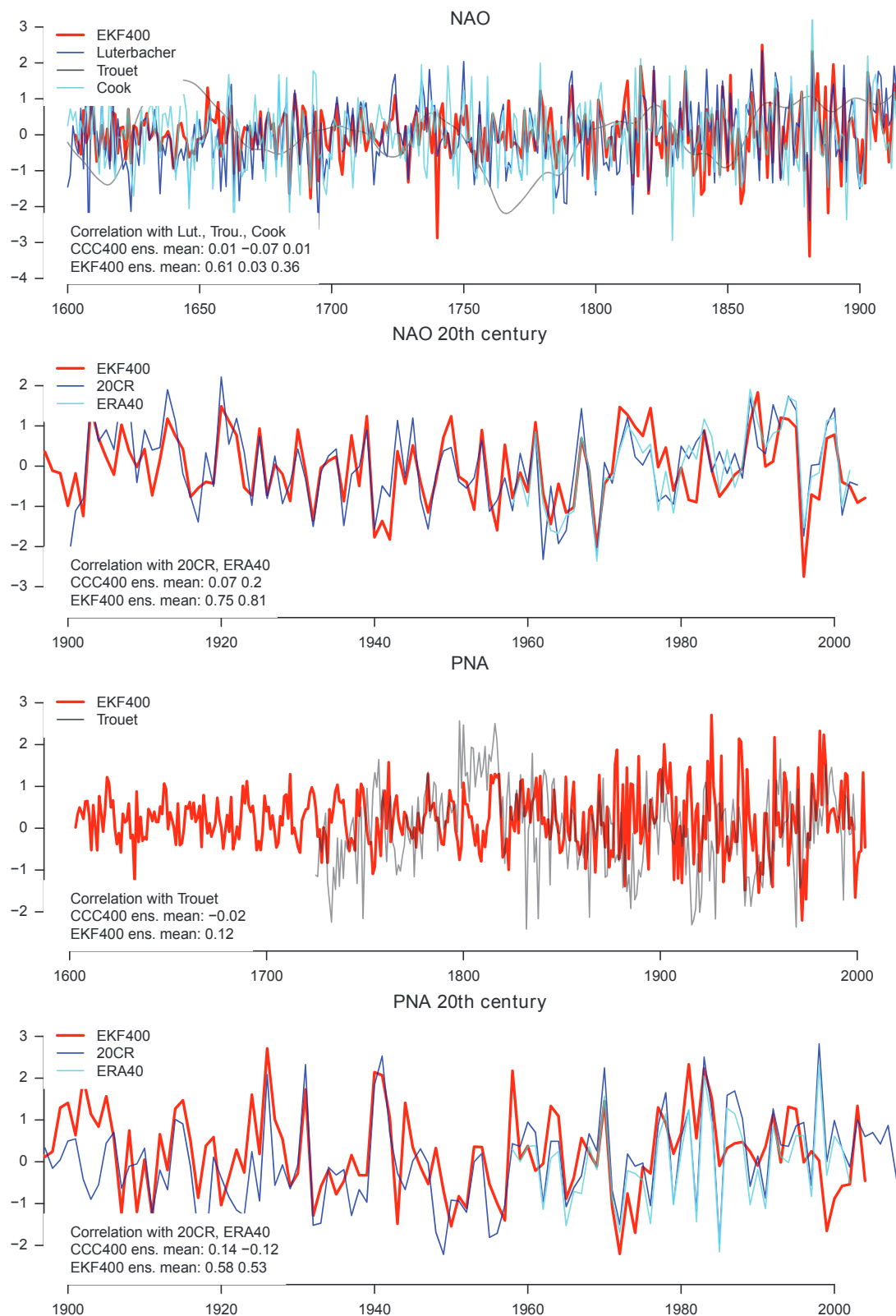


Figure 12. The availability the full atmospheric state also allows for the calculation of circulation indices. The December to March NAO/PNA index (top/bottom) of EKF400 correlates much better with the reference series during the 20th century (top) than CCC400. Values in the lower left corner indicate correlation coefficient of EKF400/CCC400 with the reference data sets for the overlapping period.

reconstructing the strength of the polar vortex, the Hadley cell, the position of the subtropical jet and the dynamic monsoon index. Out of these, we currently only see skill in the polar vortex strength at 100 hPa. For more skill at lower latitudes, our current paleodata network and covariance matrix together with the applied localization do not provide sufficient constraints.

Usage Notes

Combining direct and indirect climate observations, climate forcing time series and a climate model, this data set joins all our information including their uncertainties into the currently best estimate of monthly variability of the three-dimensional atmosphere during the past 400 years. A few thousand kilometres around assimilated data, EKF400 contains more realistic information than the unconstrained model ensemble. Hence, this new paleo-reanalysis will mainly allow for new insights around northern hemisphere land areas whereas oceanic regions closely follow the SST forcings.

The EKF400 reanalysis described in this article is available at the World Data Center for Climate at Deutsches Klimarechenzentrum (DKRZ) in Hamburg, Germany (Data Citation 1). A login is provided to everybody after sending an email to 'data@dkrz.de'.

Note, the data south of 60°S should not be used (see 'model simulation' section). In general, EKF400 has little skill in the southern hemisphere that goes beyond the original model simulations because only few observations have been assimilated.

Variability on time scales longer than 70 years stems from the model and its forcings, including SST forcings. This data set should therefore not be used to analyse multi-decadal variability of oceanic indices.

The assimilation is conducted based on anomalies. If the mean is added to receive absolute values, this may rarely result in total precipitation values below zero. Please be aware of this artefact.

Future improvements should include an extension of the paleodata network and better error assessment of the assimilation data. Largest gains in analysis skill can be expected from an improved description of the background covariance matrix, which is currently only based on the spread of the model ensemble at the assimilation time. An improved background covariance matrix would reduce the need for localization and hence allow accounting for more distant teleconnections.

References

1. Jones, P. D. *et al.* High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. *The Holocene* 19, 3–49 (2009).
2. Compo, G. P. *et al.* The Twentieth Century Reanalysis Project. *Q. J. Roy. Meteorol. Soc.* 137, 1–28 (2011).
3. Poli, P. *et al.* ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. *J. Climate* 29, 4083–4097 (2016).
4. Evensen, G. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynam.* 53, 343–367 (2003).
5. Raible, C. C., Lehner, F., González-Rouco, J. F. & Fernández-Donado, L. Changing correlation structures of the Northern Hemisphere atmospheric circulation from 1000 to 2100 AD. *Clim. Past.* 10, 537–550 (2014).
6. Bhend, J., Franke, J., Folini, D., Wild, M. & Brönnimann, S. An ensemble-based approach to climate reconstructions. *Clim. Past.* 8, 963–976 (2012).
7. Hakim, G. J. *et al.* Overview of data assimilation methods. *PAGES news* 21, 72–73 (2013).
8. Brönnimann, S. *et al.* Transient state estimation in paleoclimatology using data assimilation. *PAGES news* 21, 74–75 (2013).
9. Hakim, G. J. *et al.* The last millennium climate reanalysis project: Framework and first results. *J. Geophys. Res.-Atmos.* 121, 6745–6764 (2016).
10. Dirren, S. & Hakim, G. J. Toward the assimilation of time-averaged observations. *Geophys. Res. Lett.* 32, L04804 (2005).
11. Whitaker, J. S. & Hamill, T. M. Ensemble Data Assimilation without Perturbed Observations. *Mon. Weather Rev.* 130, 1913–1924 (2002).
12. Matsikaris, A., Widmann, M. & Jungclaus, J. H. On-line and off-line data assimilation in palaeoclimatology: a case study. *Clim. Past.* 11, 81–93 (2015).
13. Gaspari, G. & Cohn, S. E. Construction of correlation functions in two and three dimensions. *Q. J. Roy. Meteorol. Soc.* 125, 723–757 (1999).
14. Brönnimann, S. *et al.* A global historical ozone data set and prominent features of stratospheric variability prior to 1979. *Atmos. Chem. Phys.* 13, 9623–9639 (2013).
15. Franke, J., Frank, D., Raible, C. C., Esper, J. & Brönnimann, S. Spectral biases in tree-ring climate proxies. *Nat. Clim. Change* 3, 1–5 (2013).
16. Osborn, T. J. & Jones, P. D. The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. *Earth Syst. Sci. Data* 6, 61–68 (2014).
17. Dee, S. G., Steiger, N. J., Emile-Geay, J. & Hakim, G. J. On the utility of proxy system models for estimating climate states over the common era. *J. Adv. Model. Earth Sy.* 8, 1164–1179 (2016).
18. Tolwinski-Ward, S. E., Evans, M. N., Hughes, M. K. & Anchukaitis, K. J. An efficient forward model of the climate controls on interannual variation in tree-ring width. *Clim. Dynam.* 36, 2419–2439 (2011).
19. Breitenmoser, P., Brönnimann, S. & Frank, D. Forward modelling of tree-ring width and comparison with a global network of tree-ring chronologies. *Clim. Past.* 10, 437–449 (2014).
20. Wilson, R. J. S. & Luckman, B. H. Dendroclimatic reconstruction of maximum summer temperatures from upper treeline sites in Interior British Columbia, Canada. *The Holocene* 13, 851–861 (2003).
21. Grudd, H. Torneträsk tree-ring width and density AD 500–2004: a test of climatic sensitivity and a new 1500-year reconstruction of North Fennoscandian summers. *Clim. Dynam.* 31, 843–857 (2008).
22. Esper, J. *et al.* European summer temperature response to annually dated volcanic eruptions over the past nine centuries. *B. Volcanol.* 75, 1–14 (2013).
23. McCarroll, D. *et al.* A 1200-year multiproxy record of tree growth and summer temperature at the northern pine forest limit of Europe. *The Holocene* 23, 471–484 (2013).
24. Frei, C. Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. *Int. J. Climatol.* 34, 1585–1605 (2014).

25. Roeckner, E. *et al.* The atmospheric general circulation model ECHAM5 Part I: model description. Max Planck Institute for Meteorology. *Tech. rep.* 349 (2003).
26. Crowley, T. J. *et al.* Volcanism and the little ice age. *PAGES news* 16, 22–23 (2008).
27. Koch, D., Jacob, D., Tegen, I., Rind, D. & Chin, M. Tropospheric sulfur simulation and sulfate direct radiative forcing in the Goddard Institute for Space Studies general circulation model. *J. Geophys. Res.-Atmos.* 104, 23799–23822 (1999).
28. Lean, J. Evolution of the Sun's Spectral Irradiance Since the Maunder Minimum. *Geophys. Res. Lett.* 27, 2425–2428 (2000).
29. Yoshimori, M., Raible, C. C., Stocker, T. F. & Renold, M. Simulated decadal oscillations of the Atlantic meridional overturning circulation in a cold climate state. *Clim. Dynam.* 34, 101–121 (2010).
30. Pongratz, J., Reick, C., Raddatz, T. & Claussen, M. A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochem. Cycles* 22, GB3018 (2008).
31. Mann, M. E., Woodruff, J. D., Donnelly, J. P. & Zhang, Z. Atlantic hurricanes and climate over the past 1,500 years. *Nature* 460, 880–885 (2009).
32. Rayner, N. A. *et al.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.-Atmos.* 108, doi:10.1029/2002JD002670, (2003).
33. Cook, A. R. & Schaefer, J. T. The relation of El Niño-Southern Oscillation (ENSO) to winter tornado outbreaks. *Mon. Weather Rev.* 136, 3121–3137 (2008).
34. Lawrimore, J. H. *et al.* An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *J. Geophys. Res.* 116, 1–18 (2011).
35. Brohan, P., Kennedy, J. J., Harris, I., Tett, S. & Jones, P. D. Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.-Atmos.* 111, 1–21 (2006).
36. Auer, I. *et al.* HISTALP—historical instrumental climatological surface time series of the Greater Alpine Region. *Int. J. Climatol.* 27, 17–46 (2007).
37. Böhm, R. *et al.* The early instrumental bias: a solution for long central European temperature series 1760–2007. *Climatic Change* 101, 41–67 (2010).
38. Parker, D. E., Legg, T. P. & Folland, C. K. A new daily central England temperature series, 1772–1991. *Int. J. Climatol.* 12, 317–342 (1992).
39. Farrona, A. M. M., Trigo, R. M., Gallego, M. C. & Vaquero, J. M. The meteorological observations of Bento Sanches Dorta, Rio de Janeiro, Brazil: 1781–1788. *Climatic Change* 115, 579–595 (2012).
40. Trigo, R. M., Vaquero, J. M. & Stothers, R. B. Witnessing the impact of the 1783–1784 Laki eruption in the Southern Hemisphere. *Climatic Change* 99, 535–546 (2010).
41. Alcoforado, M. J., Vaquero, J. M., Trigo, R. M. & Taborda, J. P. Early Portuguese meteorological measurements (18th century). *Clim. Past* 8, 353–371 (2012).
42. Zaiki, M. *et al.* Recovery of nineteenth-century Tokyo/Osaka meteorological data in Japan. *Int. J. Climatol.* 26, 399–423 (2006).
43. Vinther, B. M., Andersen, K. K., Jones, P. D., Briffa, K. R. & Cappelen, J. Extending Greenland temperature records into the late eighteenth century. *J. Geophys. Res.* 111, 1–13 (2006).
44. Chenoweth, M. & Thistlewood, T. The 18th Century Climate of Jamaica: Derived from the Journals of Thomas Thistlewood, 1750–1786. *T. Am. Philos. Soc.* 93, 1–153 (2003).
45. Peterson, T. C. & Vose, R. S. An overview of the global historical climatology network temperature database. *B. Am. Meteorol. Soc.* 78, 2837–2849 (1997).
46. Küttel, M. *et al.* The importance of ship log data: reconstructing North Atlantic, European and Mediterranean sea level pressure fields back to 1750. *Clim. Dynam.* 34, 1115–1128 (2010).
47. Van Der Schrier, G. & Jones, P. D. Daily temperature and pressure series for Salem, Massachusetts (1786–1829). *Climatic Change* 87, 499–515 (2008).
48. Buchan, A. The meteorology of Gordon Castle. *Journal of the Scottish Meteorological Society* 5, 59–63 (1880).
49. Koninklijk Nederlands Meteorologisch Instituut. *Nederlandsch Meteorologisch Jaarboek voor 1870* (Kemink en Zoon, 1871).
50. Bartholy, J., Pongrácz, R. & Molnár, Z. Classification and analysis of past climate information based on historical documentary sources for the Carpathian basin. *Int. J. Climatol.* 24, 1759–1776 (2004).
51. Dobrovolný, P. *et al.* Monthly, seasonal and annual temperature reconstructions for Central Europe derived from documentary evidence and instrumental records since AD 1500. *Climatic Change* 101, 69–107 (2010).
52. Glaser, R. & Riemann, D. A thousand-year record of temperature variations for Germany and Central Europe based on documentary data. *J. Quaternary Sci.* 24, 437–449 (2009).
53. Przybylak, R. *et al.* Temperature changes in Poland from the 16th to the 20th centuries. *Int. J. Climatol.* 25, 773–791 (2005).
54. Pfister, C. *Weternachtersage* (Haupt-Verlag, 1999).
55. Briffa, K. R. *et al.* Tree-ring width and density data around the Northern Hemisphere: Part 1, local and regional climate signals. *The Holocene* 12, 737–757 (2002).
56. Esper, J., Büntgen, U., Frank, D. C., Pichler, T. & Nicolussi, K. In *Tree rings in archaeology, climatology and ecology* (ed. Haneca K. *et al.*) 80–85 Updating the Tyrol tree-ring dataset (Trace 5, 2007).
57. Büntgen, U., Frank, D. C., Nievergelt, D. & Esper, J. Summer temperature variations in the European Alps, AD 755–2004. *J. Climate* 19, 5606–5623 (2006).
58. Popa, I. & Kern, Z. Long-term summer temperature reconstruction inferred from tree-ring records from the Eastern Carpathians. *Clim. Dynam.* 32, 1107–1117 (2009).
59. Gunnarson, B. E., Linderholm, H. W. & Moberg, A. Improving a tree-ring reconstruction from west-central Scandinavia: 900 years of warm-season temperatures. *Clim. Dynam.* 36, 97–108 (2011).
60. Liñán, I. D. *et al.* Estimating 750 years of temperature variations and uncertainties in the Pyrenees by tree ring reconstructions and climate simulations. *Clim. Past* 8, 919–933 (2012).
61. Büntgen, U. *et al.* Filling the Eastern European gap in millennium-long temperature reconstructions. *P. Natl Acad. Sci. USA* 110, 1773–1778 (2013).
62. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 Dataset. *Int. J. Climatol.* 34, 623–642 (2014).
63. Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. & Wanner, H. European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* 303, 1499–1503 (2004).
64. Pauling, A., Luterbacher, J., Casty, C. & Wanner, H. Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation. *Clim. Dynam.* 26, 387–405 (2006).
65. Crowley, T. J., Obrochta, S. P. & Liu, J. Recent global temperature 'plateau' in the context of a new proxy reconstruction. *Earth's Future* 2, 281–294 (2014).
66. D'Arrigo, R., Wilson, R. & Jacoby, G. On the long-term context for late twentieth century warming. *J. Geophys. Res.-Atmos.* 111, 1–12 (2006).
67. Brönnimann, S. *et al.* Variability of large-scale atmospheric circulation indices for the northern hemisphere during the past 100 years. *Meteorol. Z.* 18, 379 (2009).

68. Luterbacher, J. *et al.* Extending North Atlantic oscillation reconstructions back to 1500. *Atmos. Sci. Lett.* 2, 114–124 (2001).
69. Cook, E. & D'Arrigo, R. A reconstruction of the North Atlantic Oscillation using tree-ring chronologies from North America and Europe. *The Holocene* 8, 9–17 (1998).
70. Trouet, V. *et al.* Persistent positive North Atlantic Oscillation mode dominated the medieval climate anomaly. *Science* 324, 78–80 (2009).
71. Trouet, V., Taylor, A. H., Carleton, A. M. & Skinner, C. N. Interannual variations in fire weather, fire extent, and synoptic-scale circulation patterns in northern California and Oregon. *Theor. Appl. Climatol.* 95, 349–360 (2009).
72. Brönnimann, S., Annis, J., Ewen, T. & Annis, J. An extended Pacific–North American index from upper-air historical data back to 1922. *J. Climate* 21, 1295–1308 (2008).
73. Uppala, S. M. *et al.* The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* 131, 2961–3012 (2005).
74. Cook, E. R., Briffa, K. R. & Jones, P. D. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *Int. J. Climatol.* 14, 379–402 (1994).
75. Jolliffe, I. T. & Stephenson, D. B. *Forecast Verification. A practitioner's guide in atmospheric science* (John Wiley & Sons, Ltd. 2012).
76. Ciach, G. J. & Krajewski, W. F. On the estimation of radar rainfall error variance. *Adv. water resour.* 22, 585–595 (1999).
77. Bowler, N. E. Accounting for the effect of observation errors on verification of M0GREPS. *Meteorol. Appl.* 15, 199–205 (2008).
78. Frank, D., Esper, J. E., Zorita, E. & Wilson, R. A noodle, hockey stick, and spaghetti plate: a perspective on high-resolution paleoclimatology. *Wiley Interdisciplinary Reviews: Climate Change* 1, 507–516 (2010).
79. Ahmed, M. *et al.* Continental-scale temperature variability during the past two millennia. *Nat. Geosci.* 6, 339–346 (2013).
80. Brönnimann, S. *Climatic Changes Since 1700* (Springer, 2015).
81. Raible, C. C. *et al.* Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects. *Wires Clim. Change* 7, 569–589 (2016).
82. Auchmann, R. *et al.* Extreme climate, not extreme weather: the summer of 1816 in Geneva, Switzerland. *Clim. Past* 8, 325–335 (2012).

Data Citation

1. Franke, J., Brönnimann, S., Bhend, J. & Brugnara, Y. World Data Center for Climate at Deutsches Klimarechenzentrum. http://dx.doi.org/10.1594/WDCC/EKF400_v1 (2017).

Acknowledgements

We like to thank all researchers who contributed their input data for our assimilation exercise as well as ETH Zürich and CSCS for their support in conducting the ECHAM simulations. Finally, we appreciate very much that the World Data Center for Climate at Deutsches Klimarechenzentrum is hosting the EKF400 data. The project was supported by the Swiss National Science Foundation Project REUSE and S.B. additionally by the EC FP7 project ERA-CLIM2.

Author Contributions

J.F. developed the code, conducted the reanalysis and wrote this manuscript. J.B. developed the original code, helped with mathematical optimizations and with writing. S.B. developed the idea and shaped the entire project and contributed to writing. Y.B. collected and digitized especially early instrumental data.

Additional Information

Competing interests: The authors declare no competing financial interests.

How to cite this article: Franke, J. *et al.* A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations. *Sci. Data* 4:170076 doi: 10.1038/sdata.2017.76 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files made available in this article.

© The Author(s) 2017